Psychometric testing of a Swedish version of the Apathy Evaluation Scale

Johansson, Maurits; Johansson, Per; Stomrud, Erik; Hagell, Peter; Hansson, Oskar

Published in:
Nordic Journal of Psychiatry

DOI:
10.1080/08039488.2017.1334820

2017

Document Version:
Peer reviewed version (aka post-print)

Psychometric testing of a Swedish version of the Apathy Evaluation Scale

Running title

Testing of a Swedish Apathy Evaluation Scale

Authors

Maurits Johansson MD[1,2*]. Per Johansson MD PhD[2, 3]. Erik Stomrud MD PhD[1,5]. Peter Hagell PhD[4†]. Oskar Hansson MD PhD[1,5†].

Author Affiliation

[1] Clinical Memory Research Unit, Clinical Sciences Malmö, Lund University, Malmö, Sweden.

[2]Clinical Sciences Lund/Helsingborg, Lund University, Sweden.

[3] Department of Internal Medicine, Sahlgrenska Academy, University of Gothenburg, Sweden.

[4] The PRO-CARE Group, School of Health and Society, Kristianstad University, Kristianstad, Sweden.

[5]Memory Clinic, Skåne University Hospital, Malmö, Sweden.

†Equally contributing senior authors.

*Corresponding author:

Maurits Johansson; maurits.johansson@med.lu.se

Clinical Memory Research Unit

Department of Clinical Sciences Malmö, Lund University

Simrisbanvägen 14, SE–205 02 Malmö (Sweden)

Abstract

Background: Apathy, a prevalent and clinically relevant symptom in neurodegenerative disease, is often evaluated by the instrument Apathy Evaluation Scale (AES). However, this instrument has not been translated into Swedish, halting clinical and research efforts. Furthermore, previous studies lack analyses of some basic properties such as the legitimacy of a total score, or have analyzed dimensionality by questionable methods.

Aim: To translate and psychometrically evaluate a Swedish version of the AES.

Method: The AES was translated and its psychometric properties were tested in the Swedish BioFINDER study, including cognitively well elderly, subjects with mild cognitive or parkinsonian symptoms. Psychometric analyses were conducted according to classical test theory (CTT) and aimed to resemble those performed in the English original study by Marin et al, 1991. Dimensionality was additionally analyzed on a matrix of polychoric correlations and parallel analyses.

Results: Data indicate that the Swedish AES performs satisfactorily regarding data completeness, scaling assumptions, targeting and reliability. Principal component analyses (with parallel analysis) of polychoric correlation matrices identified a single component. Convergent and discriminative validity correlations accorded with a priori expectations.

Conclusions: The study provides initial support that this Swedish AES performs similarly to the English original and exhibits acceptable psychometric properties according to CTT, including supported unidimensionality, and may be adopted for use in clinical and research settings.

Keywords

Apathy; Apathy Evaluation Scale (AES); Neurodegenerative disease; Neuropsychiatric assessment; Psychometrics.

Background

Historically, apathy has been a relatively neglected neuropsychiatric symptom, but is now gaining increased attention in neurodegenerative disease research (1). Apathy is a prevalent neuropsychiatric symptom in neurodegenerative diseases, such as Alzheimer's (AD) and Parkinson's disease (PD) (2, 3). It is already common in mild cognitive impairment (MCI), with prevalence rates varying between 11-53% in hospital-based samples (4). Few studies have addressed apathy in healthy elderly, but existing studies suggest that apathy increases with age in otherwise healthy subjects (5). Moreover, apathy has been associated with increased caregiver distress (6, 7), worse performance in activities of daily living (8) and earlier institutionalization (9), making it a clinically highly relevant condition. Apathy has also been associated with an increased risk of progression from MCI to dementia, suggesting that these patients should be more closely observed (10-13).

The fact that apathy for long has been a neglected condition could possibly be due to lack of a clear definition and diagnostic criteria. During the early 1990s, Marin defined the syndrome of apathy as "lack of motivation not attributable to disturbance of intellect, emotion, or level of consciousness", whereas it should be considered a symptom if related to one of the latter conditions (14, 15). He further proposed diagnostic criteria for the syndrome, which in short comprise a "reduction in overt behavioural, cognitive and emotional concomitants of goal-directed behaviour" (15).

These criteria have later been revised (16, 17), but consensus is still lacking.

Several instruments have been developed to assess apathy (18, 19). One of the most widespread and studied instruments is probably the Apathy Evaluation Scale (AES) developed by Marin et al, 1991 (20). It covers affective, behavioral, and cognitive aspects of apathy during the past 4 weeks. The scale was primarily developed for use among people aged above 55 and has been evaluated among, e.g., people with AD, PD, cerebrovascular disorders, major depression, schizophrenia and in healthy elderly (18, 19). The AES is available for three different rater sources including a clinician-administered semi-structured interview version (AES-C), a self-rated form (AES-S) and an informant-rated form (AES-I) for completion by proxies. The different rater sources cover the same 18 items (Figure 1) with 4 ordered responses categories (”Not at all”, ”Slightly”, ”Somewhat” and ”A lot”; scored 1-4, respectively). Following reversed scoring for 3 items, item scores are summed into a total score ranging between 18-72 (higher scores=more apathy). This requires that there are no missing item responses (15).

While the AES has been used in several countries (21-26), a version for a Swedish context is lacking, halting clinical and research efforts. Furthermore, available psychometric analyses of the AES have not addressed basic properties such as the assumptions underpinning the legitimacy of creating a simple summed total score, and assessments of scale dimensionality have been based on procedures that do not take the ordinal properties of data into consideration.

Aims

To translate the original AES-S and AES-I into Swedish versions, evaluate and compare their psychometric properties with findings from the original (15), and expand these analyses to more fully understand the appropriateness of the AES according to classical test theory (CTT) (27). This study does not comprise the AES-C.

Patients and Methods

The study was approved by the local ethical review board and was conducted according to the Declaration of Helsinki. All participants gave their written consent.

Translation of the AES

The original English source version of the AES-S and AES-I were translated into Swedish using a forward-backward translation procedure. AES was first translated into Swedish by a bilingual Swedish clinician (author). A second bilingual person (university educated language and philosophy teacher) then back-translated the Swedish translation into English. A third bilingual translator (university educated engineer) then retranslated the English versions into Swedish. These Swedish versions were then compared with the first Swedish versions and reconciled in relation to the original English versions in collaboration with a fourth bilingual clinician (author). Based on this translation, some modifications in wording were made to produce Swedish versions of the AES-S and AES-I.

Sample

Participants were recruited from the longitudinal and prospective Swedish BioFINDER study (www.BioFINDER.se) using its cohorts of cognitively healthy

elderly controls (HC) (n=226), people without dementia but with mild cognitive symptoms (MCS) (n=201), and with parkinsonian symptoms (PS) (n=88) (for flowchart see Figure 2).

HCs were recruited by random sampling from the population-based Malmö Diet and Cancer Study, Sweden (28). HCs were aged ≥60 years, had no complaints of cognitive dysfunction, did not fulfill the criteria for MCI or any other dementia, spoke Swedish well enough to not need a translator and had a Mini Mental Status Examination (MMSE) score ≥27. The MCS sample was recruited consecutively at three memory outpatient clinics in southern Sweden, and comprised people who were referred due to cognitive symptoms experienced by the patient or an informant, were aged 60-80 years, had MMSE scores >23 and who did not fulfill dementia criteria. The MCS sample was stratified into subjective cognitive decline (SCD, no measurable cognitive deficit) or MCI according to consensus criteria (29) by a senior neuropsychologist based on a neuropsychological test battery. For both the HC and the MCS cohorts significant unstable systemic illness or organ failure making it difficult to participate, significant alcohol or substance abuse or refusing lumbar puncture constituted exclusion criteria. For the HC cohort additional exclusion criteria were significant neurological or psychiatric disease, and refusing magnetic resonance imaging (MRI). MCS subjects refusing neuropsychological assessment at baseline were also excluded, as well as those with a cognitive impairment at baseline that with certainty could be explained by another condition than a neurocognitive disease. Inclusion and exclusion criteria, as well as the stratification process, have previously been described (30).

The PS sample was recruited at a South-Swedish university hospital neurology outpatient clinic. Participants were included in the study and followed repeatedly for up to 6 years if they fulfilled the diagnostic criteria of PD, PD with dementia (PDD) or Lewy body dementia (DLB) (31-33), including those with early parkinsonian symptoms. Exclusion criteria were significant unstable systemic illness and current significant alcohol or substance abuse.

Assessments and procedures

The BioFINDER cohorts are subjected to thorough clinical evaluations including cognitive testing, neurological and psychiatric assessments, cerebrospinal fluid and blood biomarker analyses and imaging techniques (www.BioFINDER.se). The AES-S and AES-I were distributed by mail to the MCS sample. The HC and PS samples completed the AES-S at a study visit (HCs at baseline and PS at baseline or follow-up) and were instructed to ask a close relative to complete the AES-I shortly after the visit and return it by mail. The Hospital Anxiety and Depression Scale (HADS) was administered at baseline or follow up visits (see above) to evaluate symptoms of depression (HADS-D) and anxiety (HADS-A) (34)**.** HADS, which is a well established assessment tool, consists of 14 items. Seven items relate to depression (HADS-D) and seven to anxiety (HADS-A). Each item is scored 0-3, the higher the score the more severe symptomatology (34, 35). A trained nurse assessed basic cognitive status according to the MMSE (36).

Analyses

The Swedish AES-S and AES-I were psychometrically tested according to CTT (27, 37-42). Except for the assessment of data completeness, people with missing item

responses were excluded in these analyses.

Data completeness, an indicator of scale acceptability among respondents, was assessed by calculating the percentage of missing data for each item and the total scores (37). Up to 10 % missing item responses has been considered acceptable (38). Scaling assumptions were tested to assess if item scores legitimately can be summed to a total score. This was addressed by calculating item mean scores, standard deviations (SDs), and corrected item-total correlations. According to CTT principles, scoring assumptions are met if item means and SDs are roughly equal and if corrected item-total correlations exceed 0.3 (27). In addition, total scores are based on the assumption that all items represent the same underpinning variable, i.e. that they are unidimensional (39). To examine the dimensionality of the two AES versions, we first aimed to replicate the approach used in the original validation of the AES (20) by employing a principal component analysis (PCA) (presumed use of varimax as rotation method and eigenvalue >1, rotation and criteria not specified by Marin et al, 1991) to identify the number of components. However, this traditional PCA approach is based on a matrix of inter-item Pearson correlations, which assumes at least interval level variables and is inappropriate for ordinal variables such as item level data (40). Furthermore, the eigenvalue >1 rule of thumb is known to be biased and tends to identify too many components, particularly when used with ordinal data (39). Hence, we also conducted a complementary PCA based on a matrix of polychoric correlations (41, 42), and determined the number of components (dimensions) based on parallel analysis (PA) (41, 42).

Targeting, which concerns whether the sample's level of (in this case) apathy matches

the range of apathy represented by the scale, was evaluated by analyses of total score distribution, skewness and floor-/ceiling effects. Floor/ceiling effects up to 20 % are generally considered acceptable. Skewness should preferably be between -1 and +1 (27).

Score reliability was assessed by Cronbach´s alpha, which should be >0.8. Due to the relatively large number of items, which inflates alpha, we also analysed the homogeneity coefficient (mean inter-item correlation), which should be >0.3 (27, 37). In addition, the standard error of measurement (SEM) was computed (SEM= SDx√1-alpha) (27).

To evaluate construct validity, we assessed Spearman's rank correlation coefficients between AES-S, AES-I and HADS-D/-A scores. Apathy and depression are concepts that are theoretically expected to be somewhat related and therefore display a moderate correlation level (between about 0.40-0.59) (43-47), whereas anxiety should exhibit a weaker correlation with apathy (between about 0.20-0.39) (47). The inter-correlation between AES-S and AES-I should be strong (>0.60) (47).

Analyses, performed using IBM SPSS Statistics version 22, FACTOR version 9.3.1 (41) and Microsoft Excel for Mac 2011 version 14.0, focus on the total sample. In table 1 and 2 the results are displayed in more detail.

Results

A total of 496 AES-S and 403 AES-I forms were collected (see figure 2). Reasons for attrition included non-consent (n=4) and lack of a suitable relative (AES-I; n=87).

Four MCS subjects could not be characterized as SCD or MCI due to incomplete neuropsychological assessments. HADS data were missing for 30 participants due to administrative errors.

Demographic and descriptive data are presented in Table 1. The sample consisted of 201 HCs (39 %), 97 people with SCD (19 %), 125 people with MCI (24 %), 71 people with PD (14 %) and 17 people with PDD/DLB (3 %).

Psychometric properties of the Swedish AES

Data completeness, scaling assumptions, targeting and reliability are presented in Table 2. Item level data completeness ranged 97.4-99.6 % for AES-S and 96.8-99.8 % for the AES-I with the lower values consistently representing items 6 and 11. Also item 15 marked out with a slightly, but acceptable, higher percentage of missing data (AES-S 1.8 % and AES-I 2,0 %). Total scores could be calculated from 92.9 % (n=461) and 91.0 % (n=367) of AES-S and AES-I forms, respectively.

In the total sample, AES-S and AES-I item mean scores ranged 1.6-2.3 and 1.8-2.2, respectively. Item SD ranged 0.8-1.1 for both forms. Corrected item-total correlations in the total sample were ≥0.37, with the lowest correlations representing item 6 for both versions. The other items displayed correlation values >0.5.

Table 3 presents PCA results. Data suitability for factor analysis for both forms were good with no zero correlations, significant Bartlett's tests and Kaiser-Meyer-Olkin statistics >0.80. Traditional Pearson based PCA suggested a two-dimensional structure for both AES versions. However, polychoric based PCAs with PA supported

unidimensionality and therefore the summation of items into a total score. The identified single component explained 61.2 % of the total variance (eigenvalue 11.1) for the AES-S and 62.8 % (eigenvalue 11.3) for the AES-I. Corresponding values for the second components were 6.5 % and 6.0 % (95^{th} percentiles for PA generated random matrices were 1.42 and 1.45, respectively).

Targeting analyses (Table 2) indicated mean total scores somewhat below the scale mid-point for both forms. However, the scale mid-point was within one SD of observed mean scores. A slight skewness, just above +1, was found in the total sample for AES-S but not the AES-I. Floor/ceiling effects were acceptable.

Reliability (Cronbach's alpha) was 0.95 for both the AES-S and AES-I (Table 2). The correlation between AES-S and AES-I for the total sample was 0.74. AES-S and HADS-D scores correlated 0.48 and AES-S and HADS-A scores correlated 0.35. Corresponding coefficients for the AES-I were 0.35 and 0.21.

Discussion

This study tested the psychometric properties of Swedish versions of the AES-S and AES-I by employing methods similar to those used for testing the original versions. Our CTT based observations indicate that the Swedish versions exhibit properties that with few exceptions are in general alignment with the original English versions. Furthermore, application of PCA methodology that account for the ordinal nature of data yielded good evidence for the unidimensionality of the AES, and thus the use of summed total scores. In addition, we found support for construct validity by demonstrating expected correlations between AES-S/AES-I and HADS-D and

HADS-A scores, as well as a strong correlation between AES-S and AES-I scores.

Scaling assumptions in the total sample provided support for legitimate summation of item scores into a total score. However, it should be pointed out that item 6 exhibited relatively low corrected item-total correlations in the separate cohorts (≤0.38). Floor and ceiling effects were generally low (<3.3%), which suggests good targeting and together with a relatively low SEM (2.7-2.9), provide evidence that the scales meet basic premises for being able to detect group differences and changes over time. This is further corroborated by reliability indices in accord with those originally presented by Marin et al., who reported Cronbach's alpha values of 0.86 for the AES-S and 0.94 for the AES-I (15). This constitutes an important aspect of the scales' usefulness in both research and clinical settings. However, empirical tests of responsiveness are required for any firm conclusions to be made in this respect.

The original study by Marin et al, 1991, using PCA, suggested that the AES-I and AES-S predominately represent single component scales, although three components per version were detected (general apathy, interest and insight/concern). The additional components accounted for low percentages of the explained variances (5-10%) (15). Also in a study by Clarke et al, 2007, using principal axis analyses with varimax rotation multidimensionality was demonstrated with the component "apathy" being the predominant one. Clarke et al, 2007, identified two components, for the AES-I "apathy" and "interest" and for the AES-S "apathy" and "other". However, the face validity of the "other" component was questioned since the items were theoretically difficult to combine, but instead represented double negative wordings

(48). One reason for such counterintuitive results regarding the dimensionality of the scale may be the use of traditional Pearson based correlation matrices and somewhat arbitrary rules for determining the number of components to be extracted (e.g., the eigenvalue >1 rule or the rule of >3 items per component). In this study we attempted to replicate the original PCA performed by Marin et al, 1991 (15), moreover we also applied an alternative approach that accounts for the ordinal nature of item level data by basing the PCAs on polychoric correlation matrices and employed PA for determining the number of components to be extracted (40-42). Accordingly, our Pearson based replicated analyses suggested a two-dimensional structure, whereas the polychoric based analyses provided clear evidence for unidimensionality. These results of unidimensionality we believe constitute an important novel finding since it more strongly suggest that the AES-S and AES-I reflect a single common psychological variable (apathy) and thus legitimise the summation of a total score.

In alignment with previous findings, our traditional PCA approach suggested a second component for both versions consisting mainly of item 15 and the double negatively worded items 6, 10 and 11. These were also the items that loaded in the third component in the study by Marin et al, 1991. (15). Despite evidence of unidimensionality in this study, these findings collectively illustrate the problems of using such negatively worded questions and moreover, possibly display the theoretical difficulty to value one's own level of understanding of one's problems (item 15). The lower (although acceptable) percentage of data completeness for these items and the lower corrected item-total correlations for item 6 may well represent additional aspects of the same problem.

In theory, the different versions of AES should be strongly correlated. In the original study, Marin et al, 1991, demonstrated a significant but moderate inter-correlation coefficient of 0.43 between AES-S and AES-I (15). In this study we report a higher correlation in support of convergent validity (0.74). Using correlations with the Zung self-rating depression and anxiety scales (49, 50), Marin et al, 1991. also suggested that the AES discriminates apathy from these conditions (AES-S vs depression 0.42, AES-S vs anxiety 0.42; corresponding figures for AES-I 0.27 and 0.23, respectively) (15). Our observations, using the HADS, are in line with these findings and support the notion that AES-S and AES-I discriminates apathy from depression and anxiety. Despite the fact that the scales by Zung were not used in this study, we consider the two scales to be of similar value, justifying the use of HADS in this setting.

The study design did not allow for assessment of test-retest reliability, to perform a multitrait-multimethod matrix as used by Marin et al, 1991, to address construct validity or to assess and evaluate the AES-C. This constitutes limitations together with the inability to have a clinician diagnose the presence or absence of apathy to gain a diagnostic golden standard for analyses of sensitivity and specificity. Although the translation process was performed systematically, the lack of native English-speaking back-translators and the involvement of authors in the process may also be seen as a limitation. Moreover, only a fraction of the participants had a diagnosis of dementia and no one had a diagnosis of AD. However, this study represents a thorough and detailed CTT based psychometric analysis of the AES-S and AES-I in a relatively large sample, which to our knowledge is the first to take account for the ordinal nature of item level data by conducting a polychoric based PCA for the assessment of dimensionality according to recommended procedures (42).

Conclusion

To conclude, our study provides initial support that this Swedish version of AES-S and AES-I works in alignment with the English original according to CTT based psychometric criteria, including evidence for scale unidimensionality, and may be adopted for use in clinical and research settings. Further studies to address test-retests reliability and to examine the AES-C are warranted. Such studies should also consider using a larger sample of subjects with dementia and using modern test theory methodologies, for example Rasch measurement theory (27), for an even more thorough assessment.

References

1.van Reekum R, Stuss DT, Ostrander L. Apathy: why care? J Neuropsychiatry Clin Neurosci. 2005;17(1):7-19.

2.Aalten P, Verhey FR, Boziki M, Bullock R, Byrne EJ, Camus V, *et al.* Neuropsychiatric syndromes in dementia. Results from the European Alzheimer Disease Consortium: part I. Dement Geriatr Cogn Disord. 2007;24(6):457-63.

3.den Brok MG, Dalen JW, van Gool WA, Moll van Charante EP, de Bie RM, Richard E. Apathy in Parkinson's disease: A systematic review and meta-analysis. Mov Disord. 2015.

4.Monastero R, Mangialasche F, Camarda C, Ercolani S, Camarda R. A systematic review of neuropsychiatric symptoms in mild cognitive impairment. J Alzheimers Dis. 2009;18(1):11-30.

5.Brodaty H, Altendorf A, Withall A, Sachdev P. Do people become more apathetic as they grow older? A longitudinal study in healthy individuals. Int Psychogeriatr. 2010;22(3):426-36.

6.Leiknes I, Tysnes OB, Aarsland D, Larsen JP. Caregiver distress associated with neuropsychiatric problems in patients with early Parkinson's disease: the Norwegian ParkWest study. Acta Neurol Scand. 2010;122(6):418-24.

7.Dauphinot V, Delphin-Combe F, Mouchoux C, Dorey A, Bathsavanis A, Makaroff Z, *et al.* Risk factors of caregiver burden among patients with Alzheimer's disease or related disorders: a cross-sectional study. J Alzheimers Dis. 2015;44(3):907-16.

8.Laatu S, Karrasch M, Martikainen K, Marttila R. Apathy is associated with activities of daily living ability in Parkinson's disease. Dement Geriatr Cogn Disord. 2013;35(5-6):249-55.

9.Bakker C, de Vugt ME, van Vliet D, Verhey FR, Pijnenburg YA, Vernooij-Dassen MJ, *et al.* Predictors of the time to institutionalization in young- versus late-onset dementia: results from the Needs in Young Onset Dementia (NeedYD) study. J Am Med Dir Assoc. 2013;14(4):248-53.

10.Robert PH, Berr C, Volteau M, Bertogliati C, Benoit M, Sarazin M, *et al.* Apathy in patients with mild cognitive impairment and the risk of developing dementia of Alzheimer's disease: a one-year follow-up study. Clin Neurol Neurosurg. 2006;108(8):733-6.

11.Richard E, Schmand B, Eikelenboom P, Yang SC, Ligthart SA, Moll van Charante EP, *et al.* Symptoms of apathy are associated with progression from mild cognitive impairment to Alzheimer's disease in non-depressed subjects. Dement Geriatr Cogn Disord. 2012;33(2-3):204-9.

12.Robert PH, Berr C, Volteau M, Bertogliati-Fileau C, Benoit M, Guerin O, *et al.* Importance of lack of interest in patients with mild cognitive impairment. Am J Geriatr Psychiatry. 2008;16(9):770-6.

13.Dujardin K, Sockeel P, Delliaux M, Destee A, Defebvre L. Apathy may herald cognitive decline and dementia in Parkinson's disease. Mov Disord. 2009;24(16):2391-7.

14.Marin RS. Differential diagnosis and classification of apathy. Am J Psychiatry. 1990;147(1):22-30.

15.Marin RS. Apathy: a neuropsychiatric syndrome. J Neuropsychiatry Clin Neurosci. 1991;3(3):243-54.

16.Starkstein SE, Petracca G, Chemerinski E, Kremer J. Syndromic validity of apathy in Alzheimer's disease. Am J Psychiatry. 2001;158(6):872-7.

17.Robert P, Onyike CU, Leentjens AF, Dujardin K, Aalten P, Starkstein S, *et al.* Proposed diagnostic criteria for apathy in Alzheimer's disease and other neuropsychiatric disorders. Eur Psychiatry. 2009;24(2):98-104.

18.Clarke DE, Ko JY, Kuhl EA, van Reekum R, Salvador R, Marin RS. Are the available apathy measures reliable and valid? A review of the psychometric evidence. J Psychosom Res. 2011;70(1):73-97.

19.Radakovic R, Harley C, Abrahams S, Starr JM. A systematic review of the validity and reliability of apathy scales in neurodegenerative conditions. Int Psychogeriatr. 2015;27(6):903-23.

20.Marin RS, Biedrzycki RC, Firinciogullari S. Reliability and validity of the Apathy Evaluation Scale. Psychiatry Res. 1991;38(2):143-62.

21.Lueken U, Seidl U, Schwarz M, Volker L, Naumann D, Mattes K, *et al.* [Psychometric properties of a German version of the Apathy Evaluation Scale]. Fortschr Neurol Psychiatr. 2006;74(12):714-22.

22.Al-Adawi S, Dorvlo AS, Burke DT, Huynh CC, Jacob L, Knight R, *et al.* Apathy and depression in cross-cultural survivors of traumatic brain injury. J Neuropsychiatry Clin Neurosci. 2004;16(4):435-42.

23.Lee SH, Wen MC, Chao CC, Chen YJ, Yen CF. Apathy in late-life depression among Taiwanese patients. Int Psychogeriatr. 2008;20(2):328-37.

24.Santangelo G, Vitale C, Trojano L, Picillo M, Moccia M, Pisano G, *et al.* Relationship between apathy and cognitive dysfunctions in de novo untreated Parkinson's disease: a prospective longitudinal study. Eur J Neurol. 2015;22(2):253-60.

25.Kasai M, Meguro K, Nakamura K. [Reliability and validity of the Japanese version of the Apathy Evaluation Scale]. Nihon Ronen Igakkai Zasshi. 2014;51(5):445-52.

26.Sagen U, Faerden A, Haug T, Melle I, Finset A, Dammen T. Are there common core features of apathy in different neuropsychiatric samples as assessed by the Apathy Evaluation Scale? Nordic journal of psychiatry. 2010;64(1):49-57.

27.Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. Health Technol Assess. 2009;13(12):iii, ix-x, 1-177.

28.Manjer J, Carlsson S, Elmstahl S, Gullberg B, Janzon L, Lindstrom M, *et al.* The Malmo Diet and Cancer Study: representativity, cancer incidence and mortality in participants and non-participants. European journal of cancer prevention : the official journal of the European Cancer Prevention Organisation (ECP). 2001;10(6):489-99.

29.Petersen RC. Mild cognitive impairment as a diagnostic entity. J Intern Med. 2004;256(3):183-94.

30.Mattsson N, Insel PS, Palmqvist S, Stomrud E, van Westen D, Minthon L, *et al.* Increased amyloidogenic APP processing in APOE varepsilon4-negative individuals with cerebral beta-amyloidosis. Nat Commun. 2016;7:10918.

31.Gelb DJ, Oliver E, Gilman S. Diagnostic criteria for Parkinson disease. Arch Neurol. 1999;56(1):33-9.

32.Emre M, Aarsland D, Brown R, Burn DJ, Duyckaerts C, Mizuno Y, *et al.* Clinical diagnostic criteria for dementia associated with Parkinson's disease. Mov Disord. 2007;22(12):1689-707; quiz 837.

33.McKeith IG, Dickson DW, Lowe J, Emre M, O'Brien JT, Feldman H, *et al.* Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. Neurology. 2005;65(12):1863-72.

34.Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand. 1983;67(6):361-70.

35.Lisspers J, Nygren A, Soderman E. Hospital Anxiety and Depression Scale (HAD): some psychometric data for a Swedish sample. Acta Psychiatr Scand. 1997;96(4):281-6.

36.Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975;12(3):189-98.

37.Ware JE, Jr., Gandek B. Methods for testing data quality, scaling assumptions, and reliability: the IQOLA Project approach. International Quality of Life Assessment. J Clin Epidemiol. 1998;51(11):945-52.

38.Saris-Baglama R, Dewey CJ, Chisholm GB, Plumb E, King J, Kosinski M, *et al*. SF Health Outcomes™ Scoring Software User's Guide. Lincoln: Quality-Metric, Inc2004.

39.Nunnally J, Bernstein, IH. Psychometric Theory. 3rd ed. New York: McGraw-Hill, Inc; 1994.

40.Carroll J. The nature of the data, or how to choose a correlation coefficient. Psychometrika. 1961;26(4):347-72.

41.Baglin J. Improving Your Exploratory Factor Analysis for Ordinal Data: A Demonstration Using FACTOR. Practical Assessment, Research & Evaluation[Internet]. 2014 [Cited 2016 Nov 1]; 19(5):[1-15 pp.]. Available from: http://pareonline.net/getvn.asp?v=19&n=5.

42.Timmerman ME, Lorenzo-Seva U. Dimensionality assessment of ordered polytomous items with parallel analysis. Psychol Methods. 2011;16(2):209-20.

43.Starkstein SE, Ingram L, Garau ML, Mizrahi R. On the overlap between apathy and depression in dementia. J Neurol Neurosurg Psychiatry. 2005;76(8):1070-4.

44.Tagariello P, Girardi P, Amore M. Depression and apathy in dementia: same syndrome or different constructs? A critical review. Arch Gerontol Geriatr. 2009;49(2):246-9.

45.Levy ML, Cummings JL, Fairbanks LA, Masterman D, Miller BL, Craig AH, *et al.* Apathy is not depression. J Neuropsychiatry Clin Neurosci. 1998;10(3):314-9.

46.Marin RS, Firinciogullari S, Biedrzycki RC. The sources of convergence between measures of apathy and depression. J Affect Disord. 1993;28(2):117-24.

47.Evans JD. Straightforward statistics for the behavioral sciences. Pacific Grove, California: Brooks/Cole Pub. Co; 1996.

48.Clarke DE, Reekum R, Simard M, Streiner DL, Freedman M, Conn D. Apathy in dementia: an examination of the psychometric properties of the apathy evaluation scale. J Neuropsychiatry Clin Neurosci. 2007;19(1):57-64.

49.Zung WW. A SELF-RATING DEPRESSION SCALE. Archives of general psychiatry. 1965;12:63-70.

50.Zung WW. A rating instrument for anxiety disorders. Psychosomatics. 1971;12(6):371-9.

Acknowledgements

The authors wish to thank the study subjects for their participation. We would also like to express our gratitude to Ida Friberg, Johanna Fredlund, Karin Nilsson, Katarina Johansson and Malin Otthén for assistance in data collection.

The study was supported by the European Research Council, the Swedish Research Council, the Strategic Research Area MultiPark at Lund University, the Crafoord Foundation, the Swedish Brain Foundation, The Swedish Alzheimer foundation, Gorthon stiftelsen, Kockska stiftelserna, Kristianstad University and the Swedish federal government under the ALF agreement. The funding sources had no role in the design and conduct of the study, in the collection, analysis, interpretation of the data, or in the preparation, review, or approval of the manuscript.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Tables and table legends

Table 1. Descriptive data according to clinical group

	N	Female sex	Age	MMSE	HADS-D[$]	HADS-A[%]
Clinical group*		N (%)	Mean (SD)	Median (q1-q3)	Median (q1-q3)	Median (q1-q3)
HC	201	125 (62.2)	75 (5)	29 (28-30)	1 (0-3)	1 (0-4)
SCD	97	54 (55.7)	70 (6)	29 (27-30)	3 (1-5)	5 (2-7)
MCI	125	60 (48.0)	71 (6)	27 (26-28)	3 (1-5)	5 (2-7)
PD	71	31 (43.7)	67 (9)	29 (27-30)	3 (2-6)	4 (2-8)
PDD/DLB	17	4 (23.5)	74 (6)	23 (20-24)	8 (5-10)	6 (4-8)
Total	511	274 (53.6)	72 (7)	29 (27-29)	2 (1-5)	3 (1-6)

Abbreviations: HC=Cognitively healthy controls, DLB=Lewy Body Disease with Dementia, F=female, HADS-A=Hospital Anxiety and Depression Scale – Anxiety, HADS-D=Hospital Anxiety and Depression Scale – Depression, MMSE=Mini Mental State Examination, N=number, MCI=Mild Cognitive Impairment, PD=Parkinson's disease, PDD=Parkinson's disease with dementia and SCD=Subjective Cognitive Decline. Missing data: *Clinical group (n=4 (due to incomplete neuropsychological assessment), $HADS-D (n=30), %HADS-A (n=30) (due to administrative errors).

Table 2. The psychometric properties of the Swedish AES-S and AES-I.

	AES-S				AES-I			
	Total (n=496)*	HC (n=199)*	MCS (n=209)*	PS (n=88)*	Total (n=367)*	HC (n=135)*	MCS (n=192)*	PS (n=76)*
Data quality*	Values	Values	Values	Values	Values	Values	Values	Values
Computable total score (%)	92.9%	95.5%	92.8%	87.5%	91.0%	93.3%	90.1%	89.5%
Item level data completeness (%)	97.4-99.6%	98.5-100%	96.7-100%	94.3-100%	96.8-99.8%	96.3-100%	95.3-99.5%	94.7-100%
Psychometric properties$								
Scaling assumptions								
Item mean score (min-max)	1.6-2.3	1.1-2.0	1.5-2.3	2.6-3.3	1.8-2.2	1.3-2.0	1.7-2.4	2.5-3.3
Item SD (min-max)	0.8-1.1	0.4-1.0	0.7-1.0	0.7-1.0	0.8-1.1	0.5-1.1	0.7-1.0	0.8-1.1
Item–total correlations% (min-max)	0.37-0.82	0.19-0.62	0.18-0.78	0.32-0.76	0.41-0.82	0.38-0.75	0.24-0.82	0.31-0.78
Targeting§								
Observed mean (SD)	34.2 (11.9)	28.0 (5.7)	32.6 (8.8)	53.3 (10.6)	36.6 (12.9)	28.7 (8.2)	36.2 (10.6)	52.3 (11.4)
Observed min-max	18–72	18-43	18-67	18-72	18-70	18-63	18-63	18-70
Floor effect (% scoring 18)	3.0%	5.3%	1.0%	2.6%	3.3%	7.1%	0.6%	2.9%
Ceiling effect (% scoring 72)	0.2%	0.0%	0.0%	1.3%	0.0%	0.0%	0.0%	0.0%
Skewness	1.1	0.3	1.0	-1.1	0.6	1.0	0.4	-0.9
Reliability								
Cronbach's alpha	0.95	0.82	0.90	0.93	0.95	0.91	0.93	0.93
Mean inter-item correlation	0.52	0.21	0.35	0.43	0.53	0.38	0.41	0.44
SEM [SDx√1-alpha]	2.7	2.4	2.72	2.83	2.9	2.5	2.9	3.0

Abbreviations: AES=Apathy Evaluation Scale, AES-I=AES Informant, AES-S=AES Self, CI=confidence interval, SD=standard deviation, SEM=standard error of measurement, HC=Cognitively healthy controls, MCI=Mild Cognitive Impairment, PS=Parkinsonian symptoms. *Calculated using complete and incorrect filled out AES-forms, no other missing data were used in the statistical analysis of data quality. $Calculated only using AES were computable total score could be obtained. %Corrected item-total correlations. §Scale mid-point=54 (min-max=18-72).

Table 3. Principal component analyses* of the Swedish AES-S and AES-I

	AES-S				AES-I			
	PearsonC$		PA-PCC%		PearsonC$		PA-PCC%	
	Item	Loading	Item	Loading	Item	Loading	Item	Loading
Component 1	1	0.689	1	0.854	1	0.685	1	0.861
	2	0.698	2	0.820	2	0.691	2	0.853
	3	0.690	3	0.831	3	0.678	3	0.833
	4	0.755	4	0.774	4	0.777	4	0.818
	5	0.738	5	0.759	5	0.797	5	0.771
	7	0.771	6	0.459	7	0.803	6	0.503
	8	0.720	7	0.784	12	0.548	7	0.800
	9	0.721	8	0.911	13	0.533	8	0.862
	12	0.674	9	0.889	16	0.670	9	0.850
	13	0.713	10	0.845	17	0.803	10	0.839
	14	0.679	11	0.610	18	0.811	11	0.636
	16	0.721	12	0.754			12	0.788
	17	0.717	13	0.776			13	0.722
	18	0.802	14	0.794			14	0.691
			15	0.760			15	0.799
			16	0.828			16	0.850
			17	0.735			17	0.806
			18	0.857			18	0.884
Eigenvalue; % variance	9.97	55.37%	11.14	61.2%	10.12	56.21%	11.3	62.81%
Component 2	6	0.773			6	0.682		
	10	0.604			8	0.599		
	11	0.722			9	0.628		
	15	0.554			10	0.649		
					11	0.720		
					14	0.472		
					15	0.661		
Eigenvalue; % variance	1.138	6.32%			1.145	6.36%		

Abbreviations: AES=Apathy Evaluation Scale, AES-I=AES informant rated, AES-S= AES self rated, PearsonC=Loading matrix Pearson correlation, PA-PCC=Loading matrix Polychoric correlation with parallel analysis. Total sample with full AES data (AES-S n=461, AES-I n=367). *Applying varimax as rotation method and using Pearson correlations and polychoric correlations with parallel analysis. $Rotated loading matrix. % Unrotated loading matrix (rotated loading matrix unachievable for Pa-PCC due to PA resulting in only 1 component). AES-S: Bartlett's statistic = 5771 (df=153), p<0.000. KMO =0.958. AES-I: Bartlett's statistic =4880 (df=153), p<0.000. KMO =0.952

Figure legends

Figure 1. The Apathy Evaluation Scale (self-rated version)

Figure 2. Flowchart of inclusion, missing data and data analyses

Abbreviations: AES-I=Apathy Evaluation Scale Informant, AES-S=Apathy Evaluation Scale Self, NP Neuropsychological. *Incorrect filled out forms are not regarded as missing data. **Incorrect filled out forms (AES-S n=35, AES-I n=36) are only integrated in statistical analyses regarding data quality.